Check 31. Smart city's context awareness using social media.pdf

by Arta Sundjaja

Submission date: 17-May-2019 03:00PM (UTC+0700)
Submission ID: 1102516845
File name: 31._Smart_city_s_context_awareness_using_social_media.pdf (323.83K)
Word count: 2228
Character count: 12060

2016 International Conference on ICT For Smart Society
Surabaya, 20-21 July 2016
ISBN : 978-1-5090-1620-4

Smart City's Context Awareness Using Social Media

Fredy Purnomo
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
fpurnomo@binus.edu

Yaya Heryadi
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
yayaheryadi@binus.edu

Ford Lumban Gaol
Doctoral of Computer Science
Bina Nusantara University
Jakarta, Indonesia
fgaol@binus.edu

Michael Yoseph Ricky
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
mricky@binus.edu

***Abstract*—The increasing number of inhabitants of a city , there will be more challenges in the management of the city. Many events that can not be controlled by either causing the slow response of the relevant institutions. Sensing the smart city through social media is offered for such a solution. Text mining is done to analyze the social media posts based on events that occurred and the emotion that follows is based on text, hashtag and geo- tagging. Methodology used is text mining approach kernel methods, particularly the support vector machine (SVM). Results are expected with this concept is the city that can listen to the aspirations and desires of the population quickly and accurately.**

***Keywords*—Smart city, text mining, context awareness, social media, SVM.**

I. INTRODUCTION

Increasing number of social media users in urban areas , leading to a new era of openness of public emotion. Most citizens are literate digital world took part in the role of citizen journalism to report events, disseminate information, to express emotions, form an opinion and to voice expectations.

"We define s Smart City as one that uses information and communications technology to make both its critical infrastructure and its components and public services more interactive, efficient and better known to its residents. In the broadest sense, a city can be regarded as 'smart' when its investments in human and social capital and in its communications infrastructure foment sustainable economic development and a high standard of living along with the wise management of natural resources by an engaged government" [1].

II. SMART CITY AND SOCIAL MEDIA

A. Smart City

Smart City is a city that is combined with computers , the data and analytics that can produce a portrait of the model and condition of the city as well as the prediction of the crucial needs of a city in order to sustain in the future.

Smart cities are all about how computers, data, and analytics, which consist of models and predictions, are being embedded into cities [9].

art City by IBM divided into 6 Characteristics [1] , [2] , the Smart Economy, Smart Mobility, Smart Government, Smart People, Smart Living and Smart Environment . Each of these characteristics has a specific role to safeguard the sustainability of a city. (See Figure 1)

Fig. 1. Six Characteristics of the Smart City Model (IBM) [2]

B. Context Awareness

People as main focus in smart city issue. A smart city aims to improve urban functions and provided services with people acting as active sensors, that includes Citizen's feelings, Perception and well-being and Emotion-aware city

There are 2 ways to do observation of the Objective and Subjective observations [3] :

Objective observations using statistical data and questionnaires without involving the emotions of the city's population, including:

- Socio-demographic statistics
- Questionnaires

- Biometric sensors

Subjective observations using direct observations or perceptions of social media involving city residents , including :

- Social Media :
  - Text : Facebook, Twitter
  - Picture : Instagram, Pinterest
  - Video : Youtube, Vine
  - Chat : Whatsapp, BBM, Telegram
  - Job : Linked In
  - Sound : Smule, Shazam,
- Related with people's perception and feelings to characterize its emotional dimension [4]

According to [5], to address the problem discussed earlier, a near real-time system has been designed that is :

- personalized: alerts are based on citizen's preferences, demographics and information in the social web
- context-aware: alerts are forwarded to the citizen depending on the citizen's current and future context
- knowledge-based: reasoning is done on the fly to infer dynamically changing knowledge that aids the system to decide whether to forward an alert to a particular citizen or not.

C. Social Media

Everyday, 40% of the world's 7 billion people go online to search, cerate, learn and engage in social networking. Billions of likes and links, tweets and thumbls, vines and videos, pics and pin and pluses.

The number of users of social media in Indonesia continues to grow in recent years. Since 2011 every year there is an increase of 30% against the new members in social media. Either old or new social media. In 2014 Indonesia became the second largest country in Fabook and the fifth largest country in Twitter.

Tabel 1 Social Media Trend in Indonesia

Preferences	Condition	Value
Media	Internet cafe	65%
	Mobile phone	48%
	Home	19 %
Internet Penetration (21%, 2nd Largest facebook nation)	Facebook	38.860.460 users
	Twitter	5.679.651 users
	Blog	5.270.658 users
	Kaskus	2.600.000 users
	Foursquare	312.000 users
Age	13 – 17 years	24%
	18 – 24 years	41%
	25 – 34 years	24%
	35- 44 years	8%
	45+ years	3%
Gender	Male	60%
	Female	40%
Twitter Usage based on days	Monday	13%
	Tuesday	14%
	Wednesday	14.10%
	Thursday	17.20 %
	Friday	13%
	Saturday	13.8%
	Sunday	14.10 %
Twitter Usage based on time	Early Morning	11%
	Morning	30%
	Afternoon	26%
	Night	33%
Time to React	Younger	First to react
	18-25	First 24 hours
	Older	Day + 1
	55+	1-2 weeks

III. METHODOLOGY

Kernel Methods are a class of algorithms for pattern analysis with a number of convenient features [6]. They can deal in a uniform way with a multitude of data types and can be used to detect many types of relations in data. Popular kernel functions include linear kernel, polynomial kernel and RBF(radial basis function) kernel [7].

Text categorization plays a crucial role in both academic and commercial platforms due to the growing demand for automatic organization of documents [7]. Kernel-based classification algorithms such as Support Vector Machines (SVM) have become highly popular in the task of text mining.

Support Vector Machine is the best kernels method to handle binary classification, noisy datasets, multiclass classification, regression and novelty detection [8].

Class Meanings Kernel [7] mainly consists of four modules: preprocessing, meaning calculation, building semantickernel, and classification. Preprocessing is the step to convertion of input documents into formatted information, this process include stemming and stopword filtering. In meaning calculation step consider the meaning values of the terms according to the classes. Building semantic modelis a process to construct proposed kernel. And then, classification step uses SVM classifier to builds a model in the training phase and this model is applied to the test examples in the test phase.(see Figure 2)

The main function of kernel method for pattern analysis are capable to classification, clustering and factor analysis [6]. Classification, is to find a function of the data that can be used to correctly assign a data item (document) to one of a finite set of categories. Support Vector Machine algorithm is the classic method from machine learning. Clustering, is partitioning the samples in coherent groups. K-means is common method in clustering. Factor Analysis, when the data is high dimensional, often the interesting information contained by the data can be explained by a number of underlying factor much smaller than this dimensionality.The best known of these is principal component analysis (PCA) and canonical correlation analysis (CCA).

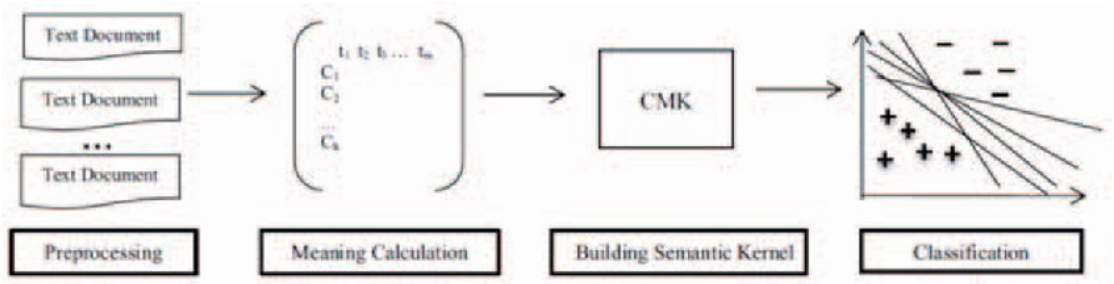

Figure 2 The Architecture of Class Meanings Kernel (CMK) System

IV. SMART CITY'S CONTEXT AWARENESS

A. City Personification

Town described as an individual who has the perspective and aware of their surroundings. A city has a population that each individual has a perspective. Perspective is meant here is the understanding of the events around and form an opinion how the pattern of view. Perspective compound of each population will be convergen form a single perspective that represents an outline of the individual perspective.

An easy way to gain perspective is through polling. However, the poll must be specifically determined at the beginning and then subsequently measured. While the events in the city is very diverse and unpredictable.

B. Sense the City through Social Media

Social media as a place to share information from the population. Any information shared, commented and responded to by the general public or a limited group. Every event can be raised issue became a hot topic. The more people who like, comment or re-share defines the success of a topic. Social media is a place for sharing of unstructured and full creativity messages of the population. If it can be quantified, it can generate the expected results.

Twitter hashtag introduced in 2007 with the hash symbol (#) followed by a word / phrase to be joined. Hastag usually placed behind the phrase that became the body of the message. This explains the hashtag type of label or metadata that is used to facilitate the search for similar topics. For example, posting on Twitter " Studying Smart City #smartcity " has hashtag #smartcity. This hashtag is used to search for similar topics with #smartcity. Can also add a hashtag corresponding equivalent word in question, for example #IOT, #smartercity, #intelligencecity or other equivalent terms.

The new social media also includes geo - tagging feature, the user can determine the location automatically when posting a message. Geotagging contains latitude and longitude coordinates of the location, and can also be the name of the place.

This research will try to make a machine that becomes a generator to generate reports of events and emotions that exist within the city based on the text input from social media by considering the location.

5W1H concept (Who, Where, When, What, Why, How) plus Impact is used to describe an event that can be categorized in a report. Measures used in developing this application is

- Social Media Text Mining (Preparation Step)
    - Twitter, Facebook, Instagram
    - Full Text Analysis
    - Hashtag Ranking
    - Body Text Analyis
    - Geotagging
- Classification
    - 5 W 1 H and Impact Analysis : Who, Where, When, What, Why, How, Impact (Activity & Emotion)
- Clustering
    - Clustering per Topic
- Factor Analysis
    - Extract the features

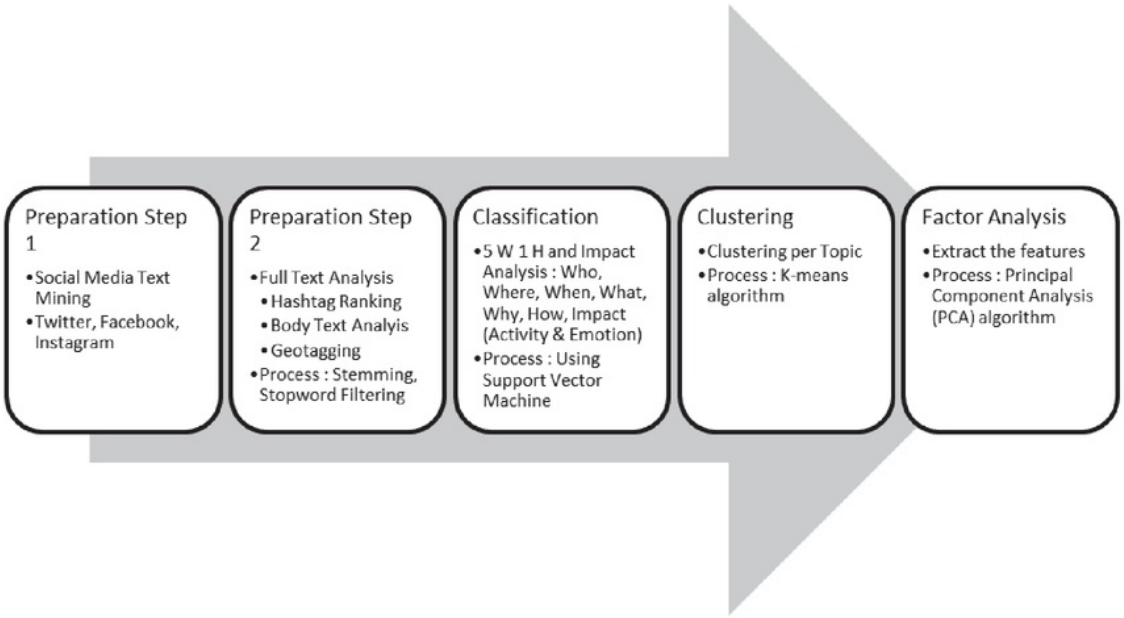

Figure 3 Text Analysis Process

Tabel 2 5W1H Analysis

IPO	Input	Process	Output
Who	Who is involved between the two parties, both of personal data, position, organization, etc	Search for words that represent people or groups and classify between the perpetrator and the victim	[Performers] [Victim]
Where	Scene including street name, building name, the name of the area, village, district, province	Search for words related to a street or place name	[Location]
When	Time events can be exact or approximate time: early morning, morning, noon, afternoon, evening and midnight.	Search for words related to temporal time	[Time]
What	What objects are involved can be	Looking for a noun as	[Object]
	everyday goods, vehicles, etc	evidence or media	
Why	Cause of the problem may be personal problems or disasters	Search for words because, using logic implications	[Cause]
How	Chronology of events	Looking for verbs associated with the way offenders	[Chronology]
Impact on Activity	Which occurs particularly due to the impact of the crowds, such as traffic jams or evacuation	Search for words related consequences caused by events	[Effect]
Impact on Emotion	Due to continued happens, especially impact the feelings and emotions	Search for words due to the related emotions and feelings	[Emotion]

V. EXPECTED RESULTS AND CONCLUSION

The expected result is the creation of a generator which will rapidly extract news from social media, without the need for people who report directly to the government. The expected result is the form of reporting to include conditions in the field that contains information about who was involved, where the scene, when it happens, what objects were involved, why did it happen, how the chronology of events and the impact on society in general in terms of both activities and emotions. In the future it could be used as an early warning for the institutions to be able to act quickly to resolve the situation that occurred in the city.

With the growing trend of social media technology, not only in the form of text and hashtag, should be considered for further research plans using video media (including text, sound and image). In general, the population has been introduced on the application of social media to share -based video, voice and text, for example Periscope. It will be a challenge for further research.

REFERENCES

[1] Priano, F. H., & Guerra, C. F. (2014). A framework for measuring smart cities. ACM, Proceedings of the 15th Annual International Conference on Digital Government Research - Dg.o '14, 44–54

[2] Giffinger, Rudolph, Christian Fertner, Hans Kramar, Robert Kalasek, Natasa Pichler-Milanović and Evert Meijers. 2007. Smart Cities: Ranking of European Medium-Sized Cities. Vienna, Austria: Centre of Regional Science (SRF), Vienna University of Technology

[3] Painho, M. (2015.). Emotion & Stress Mapping (Smart City), IEEE Xplore

[4] Guthier, B., Alharthi, R., Abaalkhail, R., & El Saddik, A. (2014). Detection and Visualization of Emotions in an Affect-Aware City. ACM DL.

[5] Banerjee, S., Mukherjee, D., & Misra, P. (2013). "What affects me?". ACM DL

[6] Bie, T. De, & Cristianini, N. (2004). Kernel methods for exploratory pattern analysis : a demonstration on text data, 1–15.

[7] Alt, B., Can, M., & Diri, B. (2015). Engineering Applications of Arti fi cial Intelligence A corpus-based semantic kernel for text classi fi cation by using meaning values of terms, 43, 54–66. http://doi.org/10.1016/j.engappai.2015.03.015

[8] Campbell, C. (1999). Chapter 7 An Introduction to Kernel Methods Kernel Methods for Classification , Regression , and Novelty Detection, 155–192.

[9] CASA - Centre for Advanced Spatial Analysis (www.mscsmartcities.org)

Check 31. Smart city's context awareness using social media.pdf

ORIGINALITY REPORT

19%	13%	17%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	openaccess.dogus.edu.tr Internet Source	5%
2	Tijl Bie. "Kernel Methods for Exploratory Pattern Analysis: A Demonstration on Text Data", Lecture Notes in Computer Science, 2004 Publication	4%
3	biblio.ugent.be Internet Source	2%
4	docplayer.net Internet Source	1%
5	Dudy Effendi, Ferra Syukri, Ahmad Fatoni Subiyanto, Rona Nandana Utdityasan. "Smart city Nusantara development through the application of Penta Helix model (A practical study to develop smart city based on local wisdom)", 2016 International Conference on ICT For Smart Society (ICISS), 2016 Publication	1%

6	Tiago H. Moreira de Oliveira, Marco Painho. "Emotion & stress mapping: Assembling an ambient geographic information-based methodology in order to understand smart cities", 2015 10th Iberian Conference on Information Systems and Technologies (CISTI), 2015 Publication	1%
7	www.broadband4europe.com Internet Source	1%
8	Snehasis Banerjee, Debnath Mukherjee, Prateep Misra. "'What affects me?'", Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication - ICUIMC '13, 2013 Publication	1%
9	Submitted to Binus University International Student Paper	1%
10	link.springer.com Internet Source	1%
11	Paulus, , and Ford Lumban Gaol. "Lung Cancer Diseases Diagnostic Asistance Using Gray Color Analysis", 2010 Second International Conference on Computational Intelligence Modelling and Simulation, 2010. Publication	1%

12	www.sustainability-seeds.org Internet Source	1%
13	Budiharto, Widodo, Welly Dwi Putra, and Yudha Pratama Putra. "Design and Analysis of Fast Driver's Fatigue Estimation and Drowsiness Detection System using Android", Journal of Software, 2013. Publication	1%
14	www.lightfc.com Internet Source	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		